

GENETIC PROGRAMMING FOR OCEAN MICROBIAL ECOLOGY AND BIODIVERSITY

John R. Moisan, NASA/Goddard Space Flight Center, Contact e-mail: John.R.Moisan@NASA.gov
Erik Wisuri, Northern Michigan University; **David Coulter**, Portland Community College;
John J. McCarthy, SGT Services, Inc.

Abstract: Present day modeling efforts to resolve upper ocean microbial ecosystem processes use coupled sets of ordinary or partial differential equations. These sets of equations, or models, which represent both ecosystem function and diversity, are subjectively developed, more or less independently, using *in situ* observations and conclusions derived from scientific literature. In the past 15 years, data assimilation (DA) techniques have become commonly used for optimizing the model parameters. Because the equations are themselves not ‘optimized’ to represent the actual ocean system under study, it has been argued that DA, targeted at parameter optimization, can only partly improve the model solutions. No objective methods existed that allowed for improving the model’s equations. This poster presentation demonstrates the application of a programming technique called “Genetic Programming” (GP) to optimize not only the set of free parameters within an ecosystem model but also the coupled set of model equations. A new GP application, termed GP COupled Differential Equations (GPCODE) is verified using a set of “twin experiments” using previously developed simple ecosystem models. The GPCODE is presently being applied to real world ecosystem observations to evolve a microbial ecosystem model for the ocean.

G: Binary tree representation allows the code to be manipulated using standard Genetic Programming principles.

- G**
- 1) Generate initial population of random models/equations.
 - 2) Calculate fitness[Sum of Squared Error (SSE)] of all individual models
 - 3) Randomly select based on fitness for: i) Asexual reproduction ii) Sexual reproduction (i.e. Tournament Selection/code crossing)
 - 4) Carry out genetic mutation
 - 5) Optimize variables using a Genetic Algorithm (GA) optimization routine
 - 6) Calculate fitness (SSE) of new individuals
 - 7) Test for completion criteria (low SSE value met?)
 - 8) No: go to 3; Yes: END Program

H: A series of ‘twin experiments’ were carried out to demonstrate the ability and robustness of this technique (GPCODE) using previously published ecosystem models (A) Simple Predator-Prey model and (B) NPZ model of Franks et al., 1986.

The road to a truly evolutionary ecosystem modeling capability, GPCODE...

A: Equations (S) need to be developed for all types of models, even for satellite algorithms.

Satellite Algorithms: S

Box Models

$$\frac{dB_i}{dt} = S_i$$

One Dimensional (Vertical) Models

$$\frac{\partial B_i}{\partial t} + (w + w_B) \frac{\partial B_i}{\partial z} - \frac{\partial}{\partial z} K_z \frac{\partial B_i}{\partial z} = S_i$$

Three Dimensional Models

$$\frac{dB_i}{dt} + (\bar{v} + \bar{v}_B) \cdot \nabla B_i - \nabla \cdot (\bar{K} \nabla B_i) = S_i$$

A

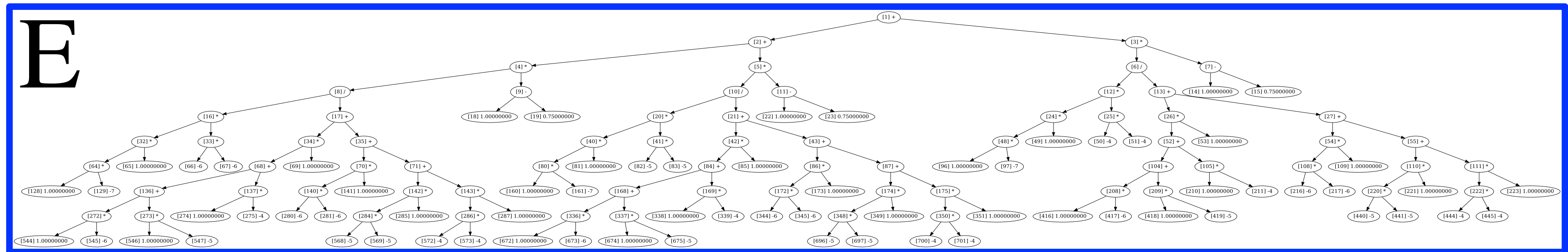
D: Binary tree representations can appear complex, even for terms in well known ecosystem models such as Fasham et al., 1990, whose binary tree for the “egestion of nitrogen into the detrital pool” equation is shown below (E).

D

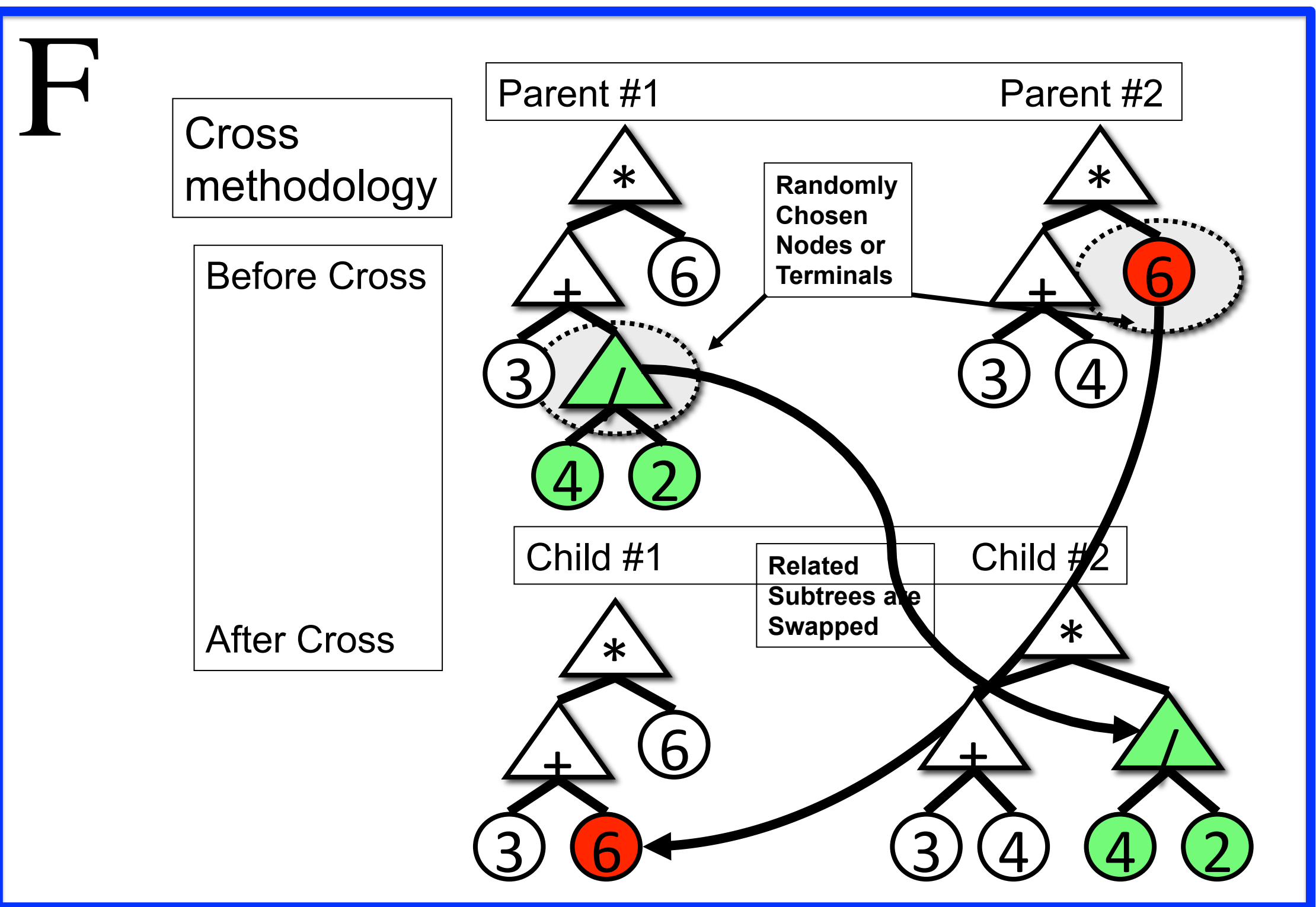
	Sink	NO ₃ ⁻ (N _n)	NH ₄ ⁺ (N _r)	DON (N _d)	DET (D)	BACT (B)	PHYTO (P)	ZOO (Z)
Source		$\frac{(m+h^*)}{M} N_o$						
NO ₃ ⁻	$\frac{(m+h^*)}{M} N_n$						$JQ_1 P$	
NH ₄ ⁺	$\frac{(m+h^*)}{M} N_r$					U_2	$JQ_2 P$	
DON	$\frac{(m+h^*)}{M} N_d$					U_1		
DET	$\frac{(m+h^*)}{M} N_o$			$\mu_4 D$				G_3
BACT	$\frac{(m+h^*)}{M} B$		$B\mu_3$					G_2
PHYTO	$\frac{(m+h^*)}{M} P$			$\gamma_1 \sigma P$	$\mu_1 P$			G_1
ZOO	$\frac{hZ}{M} + \Omega\mu_5 Z$		$(1-\Omega)\mu_5 Z$	$(1-\epsilon)\mu_2 Z$	$(1-\beta_1)G_1 + (1-\beta_2)G_2 + (1-\beta_3)G_3$			

Total RHS = Sum of Source Terms - Sum of Sink Terms; Total Possible Equations = (N)² - (N) = 56

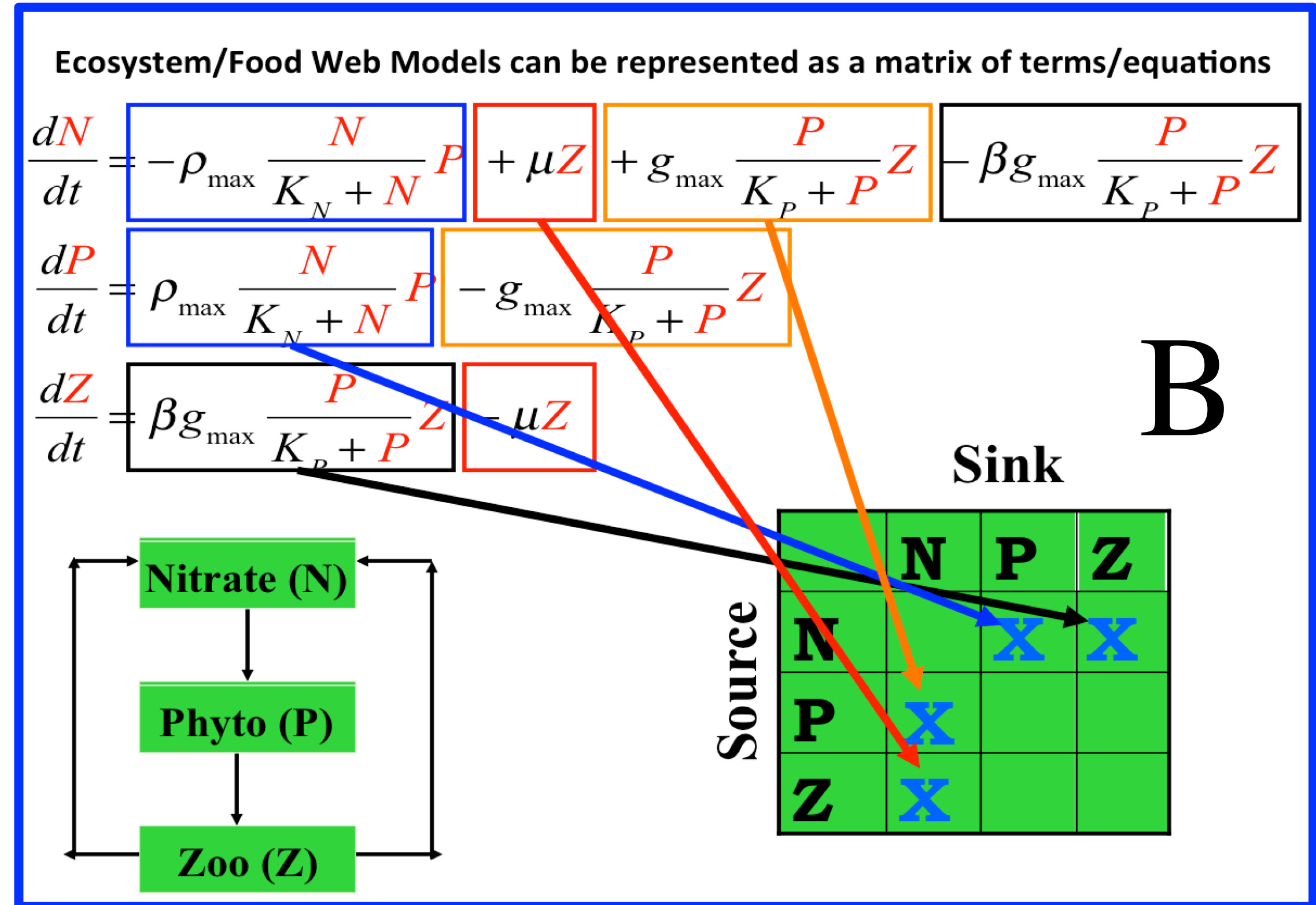
E: Binary tree representation of “egestion of nitrogen into the detrital pool” (blue box above).



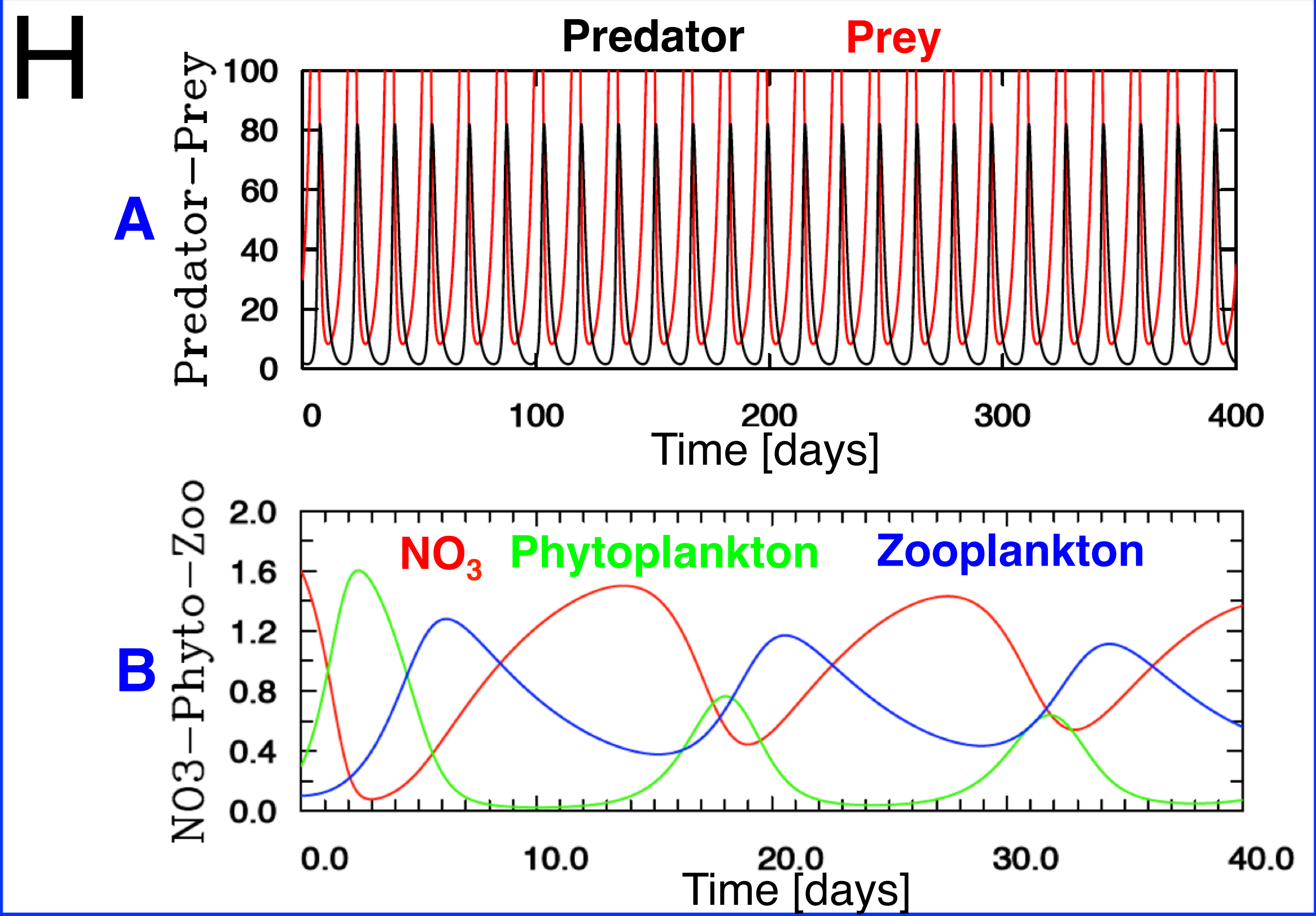
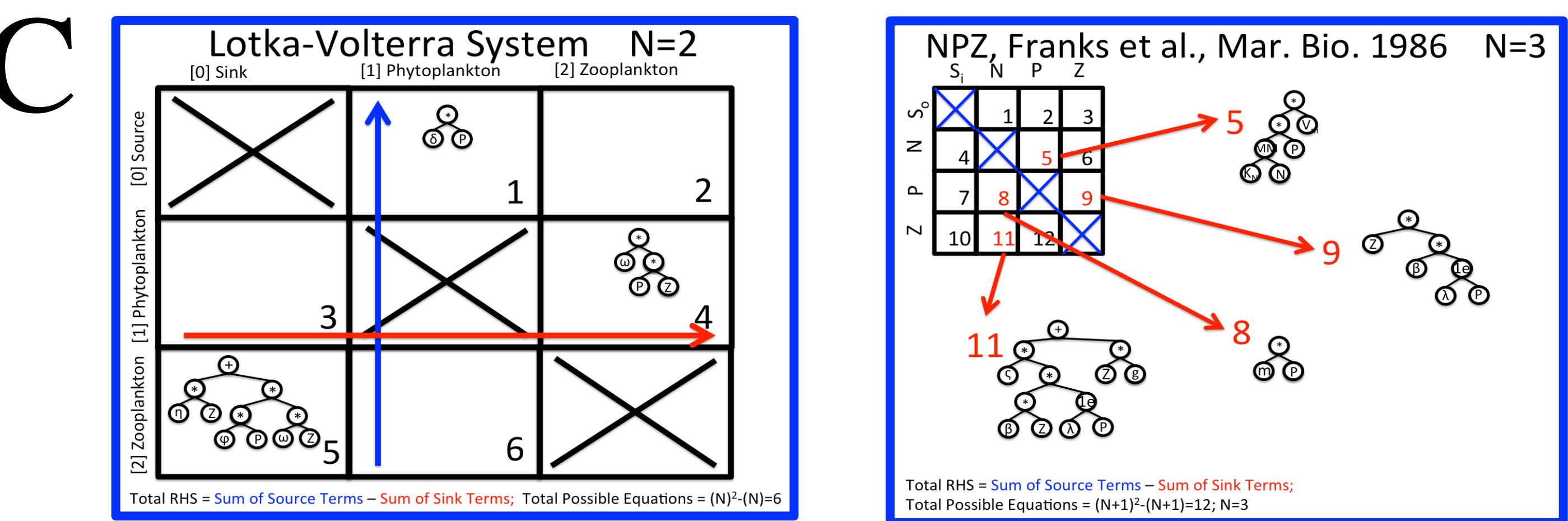
F: Binary Trees, treated as genetic information, can undergo crossing between trees.



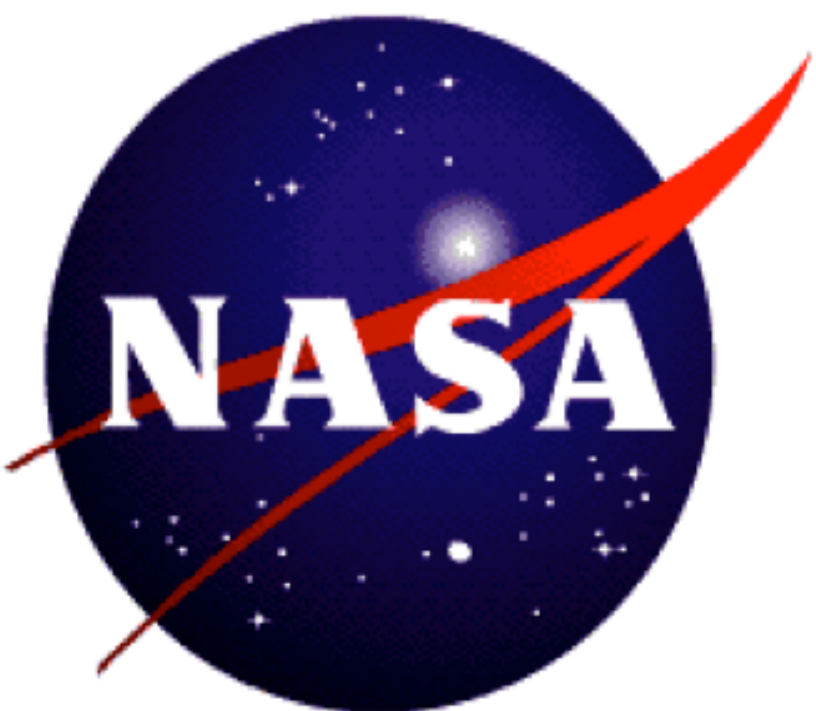
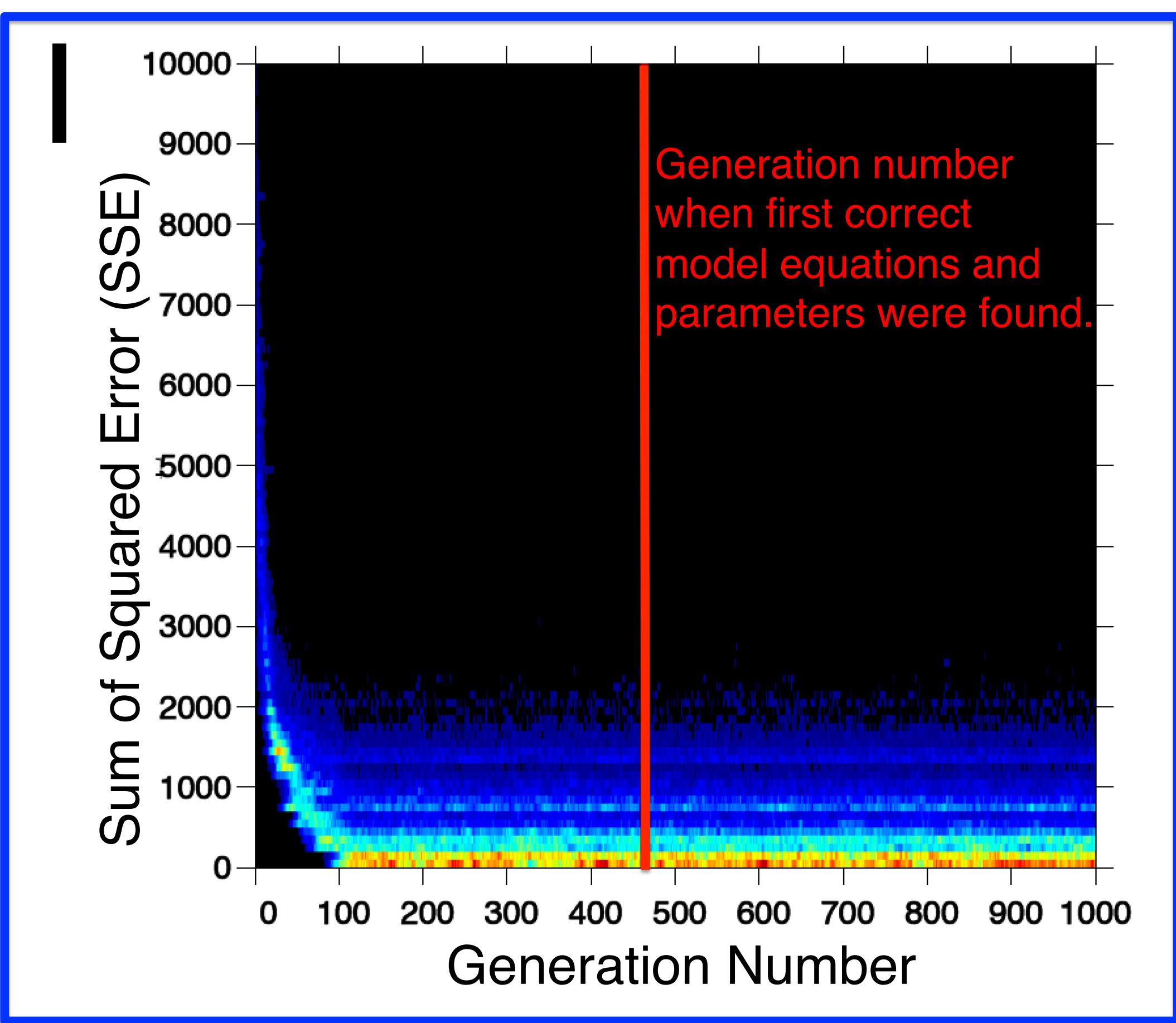
B: For ecosystem models, these equations are both coupled and can be reconfigured into an array of smaller equations.



C: Model equations can be represented as Binary-Tree Operations, (i.e. Lotka-Volterra (left) and Franks et al. 1986 NPZ (right) below).



I: The evolution of SSE histograms of the population of models shows the robustness of recovering both model equations and parameters.



GORDON AND BETTY
MOORE
FOUNDATION